STA257

Neil Montgomery | HTML is official | PDF versions good for in-class use only
Last edited: 2016-09-19 14:52

permutations and combinations

At the very least we'll need to recall (or learn!) these.

Number of ways to choose k items out of n where order matters:

$$ {}_nP_k = \begin{cases} 0 & \text{if k > n,} \\ \frac{n!}{(n-k)!} & \text{otherwise.} \end{cases} $$

and when order doesn't matter:

$$ {}_nC_k = \binom{n}{k} = \frac{n!}{k!(n-k)!} $$

Two classic examples: "The Birthday Problem" and "Lotto"

conditional probability

partial information

I'll roll a six-sided die. $S = \{1, 2, 3, 4, 5, 6\}$. Consider these events:

$$
\begin{aligned}
A &= \{2, 5\}, \\
B &= \{2, 4, 6\}, \\
C &= \{1, 2\}.
\end{aligned}
$$

So $P(A) = \frac{2}{6} = \frac{1}{3}$.

What if I peek and tell you "Actually, B occurred". What is the probabality of A given this partial information? It is $\frac{1}{3}$.

I roll the die again, peek, and tell you "Actually, C occurred". Now the probability of A is $\frac{1}{2}$.

Intuitively we used a "sample space restriction" approach.

elementary definition of conditional probability

Given B with $P(B) > 0$,

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)}$$

"The conditional probability of A given B"

The answers for the previous example coincide with the intuitive approach.

Theorem 7: For a fixed B with $P(B) > 0$, the function $P_B(A) = P(A \mid B)$ is a probability measure.

Proof: exercise.

useful expressions for calculation - I

$P(A \cap B) = P(A \mid B)P(B)$ often comes in handy.

Consider the testing for, and prevalence of, a viral infection such as HIV.

Denote by A the event "tests positive for HIV", and by B the event "is HIV positive."

For the ELISA screening test, $P(A \mid B)$ is about 0.995. The prevalence of HIV in Canada is about $P(B) = 0.00212$.

useful expressions for calculation - II

Take some event B. The sample space can be divided in two into B and B^C.

This is an example of a *partition* of S, which is generally a collection $B_1, B_2, \ldots$ of disjoint events (could be infinite) such that $\bigcup_i B_i = S$.

Theorem 8: If $B_1, B_2, \ldots$ is a partition of S with all $P(B_i) > 0$, then

$$P(A) = \sum_i P(A \mid B_i) P(B_i)$$

Proof: …

Continuing with the HIV example, suppose we also know $P(A \mid B^c) = 0.005$ ("false positive").

We can now calculate $P(A)$.

useful expressions for calculation - III

Much to my amusement, Theorem 8 gets a grandiose title: ***"THE! LAW! OF! TOTAL! PROBABILITY!!!"***

Now, in the HIV example, we also might be interested in $P(B|A)$, the chance of an HIV+ person testing positive.

A little algebra:

$$P(B \mid A) = \frac{P(B \cap A)}{P(A)} = \frac{P(A \mid B)P(B)}{P(A \mid B)P(B) + P(A \mid B^c)P(B^c)}$$

In our example this is $\frac{0.0021094}{0.0070988} = 0.2971$.

Bayes' rule

Theorem 9: If $B_1, B_2, \ldots$ is a partition of S with all $P(B_i) > 0$, then

$$P(B_i \mid A) = \frac{P(A \mid B_i)P(B_i)}{P(A)} = \frac{P(A \mid B_i)P(B_i)}{\sum_i P(A \mid B_i)P(B_i)}$$

Proof:

independence

motivation - revisit the die toss example

I'll roll a six-sided die. $S = \{1, 2, 3, 4, 5, 6\}$. Consider these events:

$$
\begin{aligned}
A &= \{2, 5\}, \\
B &= \{2, 4, 6\}
\end{aligned}
$$

So $P(A) = \frac{2}{6} = \frac{1}{3}$.

What if I peek and tell you "Actually, B occurred". What is the probabality of A given this partial information? It is $\frac{1}{3}$.

The probability of A didn't change after the new information:

$$
P(A \mid B) = \frac{P(A \cap B)}{P(B)} = P(A)
$$

definition(s) of independence

A and B are (pairwise) *independent* (notation $A \perp B$) if:

$$P(A \cap B) = P(A)P(B)$$

No requirement for $P(A)$ or $P(B)$ to be positive. In fact … see the suggested problems for Chapter 1.

$A_i, A_2, A_3, \ldots$ (possibly infinite) are (mutually) *independent* if for any finite subcollection of indices $I = \{i_1, \ldots, i_n\}$:

$$P\left(\bigcap_{i \in I} A_i\right) = \prod_{i \in I} P(A_i)$$

independence of two classes of events

Note that if $A \perp B$, then also $A \perp B^c$ and so on. Consider:

$$\mathcal{A} = \{\emptyset, A, A^c, S\}$$
$$\mathcal{B} = \{\emptyset, B, B^c, S\}$$

Classes of events $\mathcal{A}$ and $\mathcal{B}$ are *independent* all pairs of events with one chosen from each class are independent.

The suggests a concept of "independent experiments", which will be revisited.

the "any" and "all" style of examples

(Note: in probability modeling, independence is usually *assumed*.)

A subway train is removed from service if *any* of its doors are stuck open. There is a probability p of a door getting stuck open on one day of operations. A train has n doors.

What is the chance a train is removed from service due to stuck doors on one day of operations?

real valued functions with arguments that live inside sample spaces

the main focus of this course

We'll use "probability measure" throughout the course, but our main focus will be a different and equally strange object entirely.

Recall that sample space is often arbitrary and difficult or impossible to describe.

Usually we're ultimately interested in a number that is associated with the random outcome, rather than the outcome itself.

Consider a coin tossing game with $S = \{H, T\}$, which might be repeated, from which a multitude of examples can be invented.

random variable

A *random variable* is a real valued function of a sample space.

Naming convention: Roman letters near the end of the alphabet X, Y, $X_1, X_2, \ldots$.

Another strange convention - almost always omit the function's "argument".

We will never draw a picture of a random variable, or compute a derivative or an integral of one.

We will instead focus on *the* defining property of a random variable: its *distribution*.

Perversely, we will lack the math to actually define *distribution* rigorously. Informally, the *distribution* of a random variable X is the rule that assigns probabilities to values of X.